The Thesis Committee for Karl Richard Jasheway
Certifies that this is the approved version of the following thesis:

Structure Based Design of a Ricin Antidote

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

Jon D. Robertus

Marvin L. Hackert

Structure Based Design of a Ricin Antidote

by

Karl Richard Jasheway, B.S.

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

December, 2012

Abstract

Structure Based Design of a Ricin Antidote

Karl Richard Jasheway, M.A.

The University of Texas at Austin, 2012

Supervisor: Jon D. Robertus

Abstract: Ricin is a potent cytotoxin easily purified in large quantities. It presents a significant public health concern due to its potential use as a bioterrorism agent. For this reason, extensive efforts have been underway to develop antidotes against this deadly poison. The catalytic A subunit of the heterodimeric toxin has been biochemically and structurally well characterized, and is an attractive target for structure-based drug design. Aided by computer docking simulations, several ricin toxin A chain (RTA) inhibitors have been identified; the most promising leads belonging to the pterin family. To date, the most potent RTA inhibitors developed using this approach are only modest inhibitors with apparent IC_{50} values in the 10^{-4} M range, leaving significant room for improvement. This thesis discusses the development of a subset of inhibitors belonging to the pterin family in which amino acids have been utilized as building blocks. Inhibitors in this family have achieved a significant increase in potency, and have provided valuable structural information for further development.

Table of Contents

List of Tables

List of Figures

Chapter 1: Background

Ricin, from the castor plant *Ricinus communis*, is a type II holotoxin belonging to the Ribosome Inactivating Protein (RIP) superfamily [1, 2]. Type II RIPs are comprised of a catalytic A subunit, and a lectin B subunit which mediates cellular uptake. For ricin, these chains are referred to as RTA and RTB respectively. Type I RIPs consist of only the catalytic domain. Type I RIPs appear to play a role in plant antiviral defenses; they are not cytotoxic unless they can be delivered to the cytoplasm, for example by breaching the cell [3].

Ricin has received significant attention since the infamous umbrella tip assassination of Georgi Markov publically demonstrated the extreme lethality of the toxin [4, 5]. Due to its ease of extraction in large quantities from castor beans, which are processed worldwide on an industrial scale, there is a real threat of ricin being used as a biowarfare agent. It is therefore important to develop an antidote for the deadly toxin as a defense against such an attack.

The use of structure based drug design is an attractive approach for the development of small molecule inhibitors for the treatment of ricin intoxication. The use of X-ray crystallography and/or NMR spectroscopy to obtain structural information detailing the interaction between an inhibitor and its target macromolecule is the cornerstone of structure based drug design. The X-ray structures of ricin and RTA are known [6, 7, 8], and complexes with substrate analogs have revealed key features of the RTA active site [9, 10]. When the macromolecular target structure is known, medicinal chemists can rationally develop synthetic derivatives of an existing inhibitor to improve potency by creating more favorable binding interactions with the target. The work

presented in this thesis focuses on the use of this approach in the development of inhibitors targeting the catalytic A subunit of ricin.

Chapter 2: Ricin Structure and Action

2.1 – X-ray structure of ricin

The X-ray structure of the ricin holotoxin was initially solved to 2.8 Å resolution [7] and later refined at 2.5 Å [8], allowing the molecular description of the individual protein chains [8, 11]. The cloned A chain was later crystallized and solved in two different space groups at 2.1 Å resolution [20] and 1.8 Å respectively [12]. The X-ray structures allow an analysis of the suitability of each chain as a drug design target.

2.2 – RTB not suitable for drug design

RTB would be a sound choice for inhibitor design if small molecules could be made that would bind tightly to it and prevent the holotoxin from entering cells. The analysis of the X-ray structure showed that the B chain of ricin is composed of two domains, which are each composed of three related subdomains. Only one subdomain of each domain binds galactosides, and these two binding sites are over 50 Å apart (Figure 1), on opposite ends of the protein [13, 14]. The binding sites individually exhibit only weak binding to galactosides [16] with Kd values in the millimolar range. This weak binding at each site is biologically tolerable because the two widely separated sites contribute independently to the free energy of binding, and because the target cell surface is literally covered with galactosides [17]. This is not useful for inhibitor design, however. The RTB galactose binding pockets are small and make only weak interactions with galactose [13, 14]. Designing effective ligands to the shallow, polar galactose binding sites is difficult [15, 23], and the two sites are also too far apart for a small molecule to bind both sites simultaneously.

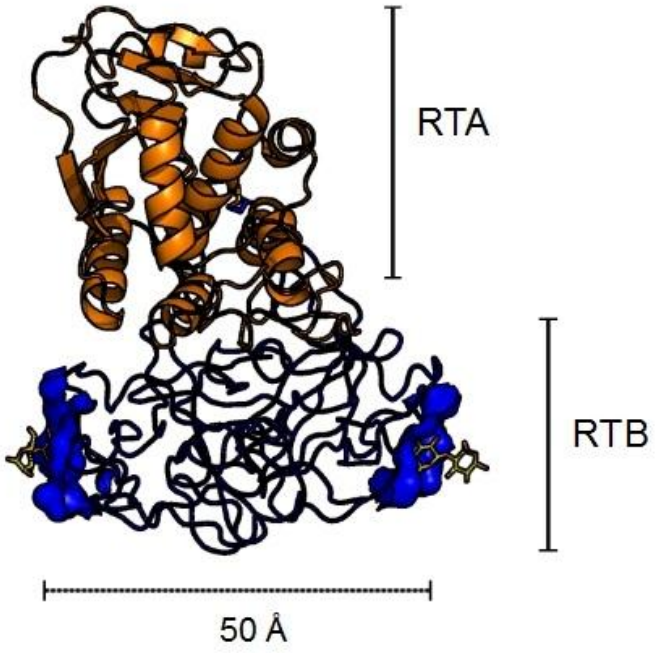

Figure 1: Xray structure of ricin holotoxin (PDB ID: 2AAI). The A chain is shown in orange and the B chain is colored blue, with galactoside binding sites shown as surfaces.

2.2 – RTA is a reasonable target for inhibitor design

Ricin Toxin A chain chemically inactivates the eukaryotic ribosome by hydrolysis of a single adenine base (A4324) on the sarcin-ricin loop (SRL) of the 28S rRNA of the large subunit [18, 19]. Ricin shows a Km for ribosomes around 1 μM, and a kcat of around 1500 min-1, depending on the ribosome species [19, 20, 21]. The catalytic efficiency of this hydrolysis reaction, kcat/Km, is near the diffusion limit. This means that ricin has evolved to enzymatic perfection for this specific ribosome inhibiting reaction. In contrast, ricin attacks naked RNA at a rate about 104-105 times more slowly [19], and only at nonphysiological (acidic) pH [22], suggesting this activity is essentially a nonspecific side reaction of its biological function [13].

The micromolar Km for ribosomes is indicative of the tight binding affinity that RTA has for its natural substrate. It is useful in structure-based inhibitor design to understand the chemical nature of that binding. The heart of substrate binding is the accommodation of the target adenine base in a "specificity" pocket in the RTA active site. The nature of this interaction was observed in a complex with formycin monophosphate (FMP), a non hydrolyzable analog of AMP [9]. Crystallographic studies of RTA showed that in the absence of substrate, the RTA specificity pocket was "closed"; that is, the side chain of Tyr 80 rotated to block its entrance [14]. However, in the presence of a substrate analog, RTA adopts an "open" conformation in which Tyrosine 80 moves to accommodate the substrate, forming a π-stacking network with the adenine base and Tyrosine 123 (Figure 1). In addition to the π-stacking interactions, the substrate forms six hydrogen bonds with the binding pocket, conferring specificity for the adenine base. Successful design of potent inhibitors of RTA is expected to require that both the π stacking and hydrogen bonding interactions be retained.

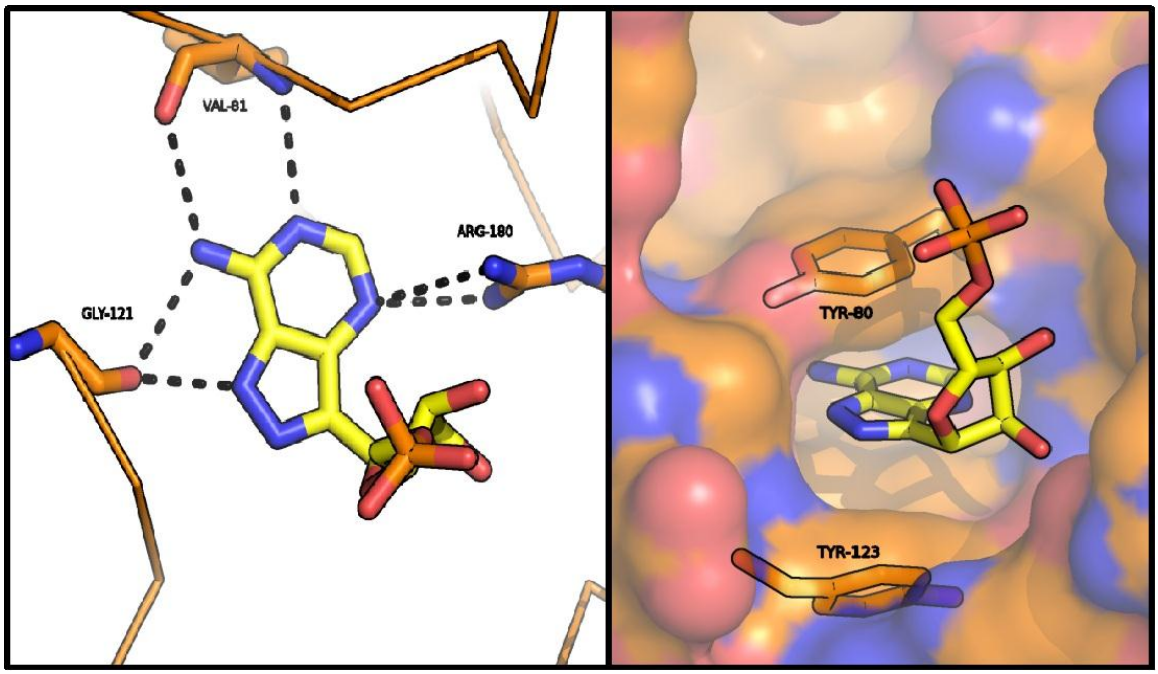

Figure 2: Structure of RTA complexed with substrate analog, formycin 5'-monophosphate (PDB ID: 1FMP): The binding of the AMP analog formycin 5' monophosphate to the active site of RTA is shown above. The aromatic ring analogous to the adenine base of the natural substrate is sandwiched between Tyrosine 80 and 123 (right), and makes six hydrogen bonds, shown as dashed black lines (left), within the active site.

The active site of RTA can be described as having two binding pockets when it is in the open conformation: the primary adenine specificity pocket and a slightly larger secondary pocket. These two pockets are separated by the side chain of Tyrosine 80. The second pocket was proposed, based on model building, to accommodate a guanine base from the invariant GAGA ribosomal target sequence [9]; this has been confirmed by the X-ray structure of an RTA complex with a locked cyclic nucleotide [10]. The guanine base forms an aromatic stack with Tyrosine 80, and thereby forms an extended stack of Tyr 123, the adenine in the specificity pocket, Tyr 80 and the guanine base. However, the binding of guanine appears to be weak as we have been unable to soak the free base,

nucleoside or nucleotide into that site. Its observed binding in the cyclic tetranucleotide is speculated to be driven by the conformational rigidity of that ligand which reduces configurational entropy of binding. Our efforts to construct small dinucleotide substrate analogs that bind to both pockets have been unsuccessful, illustrating the importance of conformational rigidity required for occupation of the second pocket.

Chapter 3: Ricin Inhibitors

3.1 – Ricin inhibitors show poor RTA inhibition.

Recently, ricin inhibitors have been identified through classical high throughput screening strategies using large libraries of compounds to protect cultured cells from intoxication [24]. However, ricin intoxication is a complex process involving cell uptake, trafficking to the ER, release to the cytoplasm, and ribosome inactivation. The cell based assay does not identify which process or protein [16] an anti-ricin compound is acting on. Another cell-based screen identified an anti-ricin compound that acts, not by inhibiting RTA action, but by disrupting cell trafficking [25]. Without a clear understanding of the macromolecular target being inhibited, it is difficult to rationally improve upon the initial high throughput hits. Of those inhibitors identified in the cell-based assay, only a small percentage showed anti-RTA activity in cell-free systems. Those compounds that did inhibit RTA were only weak inhibitors, and efforts to obtain a complex structure of these compounds with RTA have been fruitless.

3.2 – Pteroic Acid as a lead compound for RTA inhibitor design

To date, compounds that inhibit RTA action have been mainly discovered by virtual screening and structure-based design [27, 29, 30]. Such compounds have been identified using computer simulated docking of large virtual compound libraries to a model of the open form of RTA and confirming their anti-RTA activity using cell-free translation assays [27, 28, 29]. Many of these were successfully soaked or co-crystallized with RTA, and their X-ray structures revealed binding to the adenine specificity pocket. The inhibitors that yielded complex structures had numerous structural similarities, the most important of which being an exocyclic amine that donates two hydrogen bonds to the backbone carbonyls of Valine 81 and Glycine 121. Another important characteristic is

aromaticity, which is necessary for the stacking interaction with the two tyrosine side chains.

The first RTA inhibitor identified from virtual screening was pteroic acid, PTA, which had an apparent IC_{50} of 600 μM [29]. The crystal structure of the RTA-pteroic acid complex (Figure 2) reveals that the pterin group binds in the adenine specificity pocket, making six hydrogen bonds. Also, the benzoic acid moiety is in close proximity to the secondary pocket, and makes an additional hydrogen bond. Unfortunately, efforts to improve the inhibitory activity of pteroic acid by attaching pendants at the benzoic acid group were unsuccessful due to synthetic restrictions.

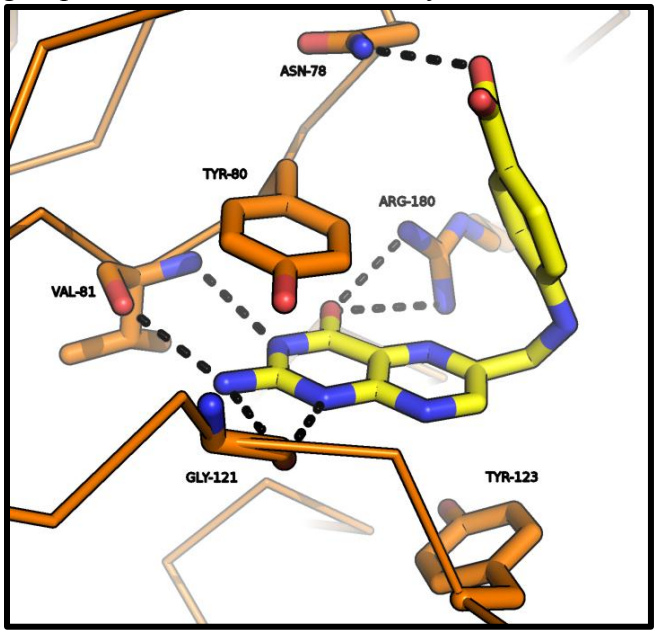

Figure 3: The crystal structure of the complex of RTA with pteroic acid is shown (PDB ID: 1BR6). Hydrogen bonds are depicted as dashed black lines.

Chapter 4: Design of Novel Pterin-based Inhibitors

Although the pterins have limited solubility, their ability to interact with a number of specific groups in the RTA active site makes them an attractive drug design platform. Based on the structural data for pterin binding in the specificity pocket, it is apparent that synthetic extensions can be made from pterin at positions 6 and 7, which could, in principle, reach toward the second pocket on the enzyme surface; most naturally occurring pterins such as pteroic acid, neopterin, and folate are substituted at the 6 position [31, 32, 33, 34].

4.1 – 7-substituted pterins are more effective RTA inhibitors

Synthetic chemists from Eric Anslyn's lab made early efforts to derivatize at the 6 and 7 positions via the addition of a carboxylic acid group. This was intended to enhance the solubility of the compounds in aqueous media, as well as adding an attachment point for nucleophiles, i.e. primary amines. Upon testing both 6 and 7 carboxy pterin (7CP) for RTA inhibition, the results overwhelmingly favored the 7 isomer. The apparent IC_{50} of 7CP was 240 μM, a nearly 3-fold improvement over pteroic acid, while 6CP had no measureable inhibition [35, 36]. It was therefore decided that synthetic optimization of the pterins would proceed exclusively off of the 7CP platform.

4.2 – Amino acids used as building blocks for 7CP derivatives

Various 7CP derivatives were synthesized using simple primary amines to attach pendants via nucleophilic attack on the carboxyl group. This gave mixed results, most compounds being weaker inhibitors than the parent compound with the pendant often being too labile to be resolved in the complex structures. In an effort to cost-effectively create a diverse collection of longer and more complex pendants, a systematic approach

was adopted in which amino acids and short peptide chains were conjugated to 7CP. The resulting compounds bear long yet rotationally limited extensions with the potential to reach the secondary binding pocket of RTA and make some intermolecular contacts between the two pockets to reinforce the binding conformation of the ligand. All compounds presented were synthesized by Jeff Pruet (7CP) and Ryota Saito (RS-series) from the Anslyn lab.

Chapter 5: Methods

5.1 – Firefly Luciferase Assay

RTA inhibition was measured using an in vitro translation assay. Firefly luciferase mRNA was translated in a cell free rabbit reticulocyte lysate system. The compounds to be tested were solubilized in 0.05 N KOH prior to their inclusion in the assay. The assay was run in the presence of BSA at concentrations in excess of that of RTA in order to neutralize compounds which non-specifically aggregate and sequester proteins, which is a common problem in assays like this one [37]. Prior to the start of the assay, the compounds are mixed with RTA and BSA and allowed to sit at room temperature for a 30 minute pre-incubation. Translation is initiated by the addition of rabbit reticulocyte lysate and translation mix (amino acids, RNasin ribonuclease inhibitor, luciferase mRNA). The reaction mixture is incubated at for 90 minutes at 30 °C, after which the reaction is quenched by freezing at -20 °C. After thawing at room temperature, the reaction mixtures are mixed with the Luciferase Substrate Reagent (Promega), and their luminescence measured on a Perkin Elmer Envision luminometer (Waltham MA).

For each concentration of compound that is to be tested, a reaction is run both in the presence and absence of RTA. The translation inhibition due to RTA, independent of effects caused by the compound, is quantified by the difference in luminescence between the two reactions. This value divided by the control translation inhibition calculated for the reactions without any test compound to yield percent RTA activity. These values are plotted and fitted to a hyperbolic decay equation using least squares regression analysis to derive the IC_{50}.

5.2 – Solution of complex structures

RTA crystallizes in two space groups: monoclinic (P21) and tetragonal (P41212). Monoclinic crystals grow in 0.8% PEG 8000 75 mM Tris-HCl pH 8.9 and require seeding for nucleation. Tetragonal crystals grow in 1.2 M ammonium sulfate 100 mM sodium malonate pH 6.0 and nucleate spontaneously. After the 7CP structure was solved using the monoclinic crystal form, the tetragonal crystals were used for subsequent experiments because of their more consistent tolerance of changing conditions such as soaking pH and cryoprotection. All crystals were grown and stored at 4 °C.

Compounds were soaked into RTA crystals by transferring the crystals into drops of mother liquor containing 1 mM of compound. Crystals were allowed to soak for no less than 24 hours prior to data collection. Data were collected with synchrotron radiation from beam line 5.0.2 of the Advanced Light Source (ALS) in Berkley, CA. Data sets were indexed, integrated, and merged using HKL2000 [38]. Phasing was done via molecular replacement using Molrep from the CCP4 software suite [39]. The search model used was a modified version of the RTA structure (PDB ID: 1RTC) in which TYR 80 was replaced with alanine. This was a precautionary measure to avoid phasing errors due to the position of the tyrosine side chain in the closed form model overlapping with the position of the ligand occupying the active site. Following refinement using REFMAC5 [40] and model editing using COOT, compounds were modeled into the unoccupied electron density in the binding pocket. Restraints necessary for modeling and refinement of the compound coordinates were generated using the PRODRG server [41]. Refinement and processing statistics are shown in Tables 1 through 3 below.

Table 1. Data collection and refinement statistics for 7CP and single amino acid derivative complex structures.

	7CP	RS2-021-1	RS2-022-1
Resolution range (Ã…)	50 - 1.29 (1.31 - 1.29)	50 - 1.7 (1.73 - 1.7)	50 - 1.44 (1.46 - 1.44)
Space group	P 1 21 1	P 41 21 2	P 41 21 2
Unit cell	42.662 67.567 49.41 90 112.78 90	67.947 67.947 140.783 90 90 90	67.805 67.805 140.813 90 90 90
Total reflections	227142	520645	818515
Unique reflections	61936	36849	59509
Multiplicity	3.7 (3.6)	14.1 (14.3)	13.6 (10.9)
Completeness (%)	95.31 (91.74)	99.75 (97.49)	98.94 (99.59)
Mean I/sigma(I)	18.45 (3.46)	20.82 (3.27)	21.06 (2.84)
Wilson B-factor	14.13	22.26	17.72
R-sym	0.042 (0.276)	0.087 (0.766)	0.079 (0.844)
R-factor	0.2069 (0.2281)	0.2132 (0.2403)	0.2051 (0.2731)
R-free	0.2327 (0.2481)	0.2523 (0.2989)	0.2348 (0.2918)
Number of atoms	2171	2327	2549
protein	2041	2078	2122
ligands	15	39	61
water	115	210	366
Protein residues	258	263	263
RMS(bonds)	0.028	0.022	0.043
RMS(angles)	2.18	2.48	3.27
Ramachandran favored (%)	99	98	98
Average B-factor	16.00	25.20	22.60
protein	15.70	24.90	20.50
solvent	20.10	29.00	33.60

Statistics for the highest-resolution shell are shown in parentheses.

Table 2. Data collection and refinement statistics for 7CP-dipeptide complex structures.

	RS2-058-1	RS2-136-1	RS2-150-1
Resolution range (Ã…)	50 - 1.52 (1.55 - 1.52)	50 - 1.58 (1.61 - 1.58)	50 - 1.75 (1.78 - 1.75)
Space group	P 41 21 2	P 41 21 2	P 41 21 2
Unit cell	67.999 67.999 140.692 90 90 90	68.009 68.009 141.023 90 90 90	67.762 67.762 140.657 90 90 90
Total reflections	707566	667334	277750
Unique reflections	51044	46156	31069
Multiplicity	13.8 (13.3)	12.2 (3.9)	8.9 (9.0)
Completeness (%)	99.23 (99.82)	99.99 (100.00)	91.80 (99.34)
Mean I/sigma(I)	19.09 (2.86)	22.39 (2.17)	15.11 (2.52)
Wilson B-factor	18.84	19.21	21.28
R-sym	0.098 (0)	0.098 (0)	0.087 (0)
R-factor	0.2011 (0.2426)	0.2121 (0.2510)	0.2372 (0.3064)
R-free	0.2304 (0.2949)	0.2391 (0.2952)	0.2679 (0.3519)
Number of atoms	2510	2375	2237
protein	2083	2110	2117
ligands	30	26	39
water	397	239	115
Protein residues	263	263	263
RMS(bonds)	0.028	0.011	0.039
RMS(angles)	2.19	1.38	1.99
Ramachandran favored (%)	98	98	98
Ramachandran outliers (%)	0.38	0.38	0.38
Average B-factor	23.20	22.10	23.60
protein	21.10	20.90	23.30
solvent	34.50	29.90	28.00

Statistics for the highest-resolution shell are shown in parentheses.

Table 3. Data collection and refinement statistics for 7CP-tripeptide complex structures.

	RS2-123-2	RS2-149-1
Resolution range (Ã…)	50 - 1.7 (1.73 - 1.7)	50 - 1.75 (1.78 - 1.75)
Space group	P 41 21 2	P 41 21 2
Unit cell	67.723 67.723 140.592 90 90 90	67.689 67.689 140.72 90 90 90
Total reflections	416685	464449
Unique reflections	33719	31660
Multiplicity	12.3 (12.4)	14.6 (14.9)
Completeness (%)	91.34 (99.39)	93.52 (90.03)
Mean I/sigma(I)	12.07 (2.69)	18.75 (3.22)
Wilson B-factor	24.76	23.72
R-sym	0.091 (0)	0.096 (0)
R-factor	0.2159 (0.2595)	0.2215 (0.2627)
R-free	0.2584 (0.2982)	0.2589 (0.3358)
Number of atoms	2347	2242
protein	2090	2078
ligands	65	78
water	192	86
Protein residues	263	263
RMS(bonds)	0.037	0.023
RMS(angles)	2.94	2.35
Ramachandran favored (%)	98	98
Ramachandran outliers (%)	0.76	0
Average B-factor	25.60	24.00
protein	24.30	23.70
solvent	34.20	24.70

Statistics for the highest-resolution shell are shown in parentheses.

Chapter 6: Results

6.1 – 7-carboxy pterin

The compound, 7 – carboxy pterin (7CP), was tested in the luciferase assay in triplicate, and had an IC_{50} of 240 μM (Figure 4A). The crystal structure of 7CP in complex with RTA, solved in the P21 space group to a resolution of 1.29 Å, is shown in Figure 4B. As seen in previous structures of RTA in complex with pterins, the 2-amino group of the pterin ring donates one hydrogen bond each to the carbonyl oxygen atoms of Val81 and Gly121, and the 4-oxo and N5 atoms accept hydrogen bonds from Arg180. The tautomeric form of the pterin with a proton on N1 is favored so that N1 can donate a hydrogen bond to the carbonyl oxygen of Gly121 and N3 can accept a hydrogen bond from the amido N of Val81. An additional hydrogen bond, which was not seen in the previous 6-substituted pterins, is made between the amido N of Tyr123 and the carbonyl oxygen of the 7-carboxy group. This extra hydrogen bond is retained in all subsequent amide derivatives based on 7CP owes to the general increase in potency of this series of pterin-based inhibitors over previous ones.

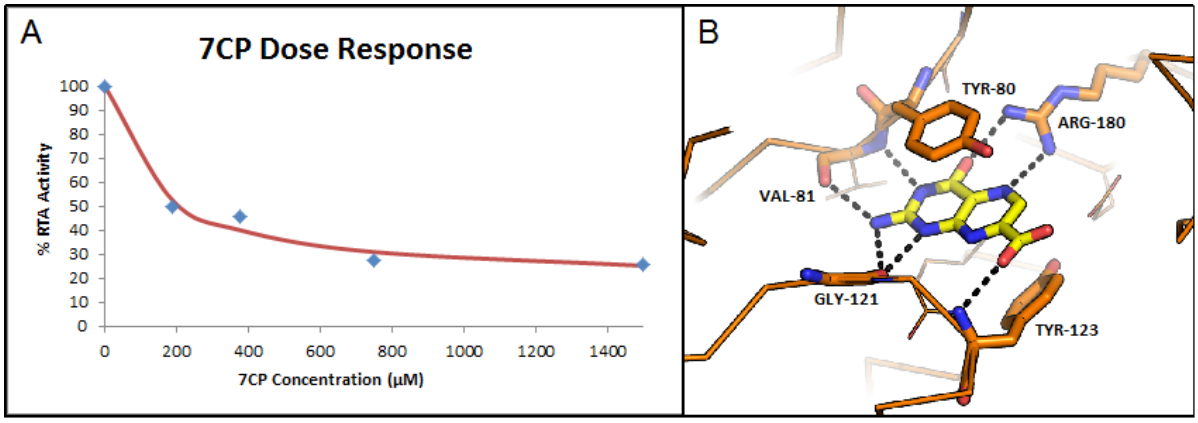

Figure 4: **A.** Dose response of 7-carboxy-pterin averaged from six repetitions of the firefly luciferase assay. The IC_{50} was calculated as 240 µM. **B.** Complex crystal structure of 7CP (yellow) and RTA (orange). The compound forms 7 hydrogen bonds, shown as black dashed lines, with RTA.

6.2 – Single amino acid derivatives

The two single amino acid 7CP derivative compounds presented in this thesis are 7CP-alanine (RS2-021-1) and 7CP-serine (RS2-022-1). Having an apparent IC_{50} value of 350 µM, the alanine derivative shows a slight decrease in potency from its parent compound. The crystal structure, solved in the P41212 space group at a resolution of 1.7 Å, shown in Figure 5B, shows that the alanine side chain makes only a minimal contribution to the binding affinity via a weak hydrophobic contact with the alpha carbon of Asn 209. The addition of two rotatable bonds and only a weak hydrophobic contact are consistent with the compound's decreased potency.

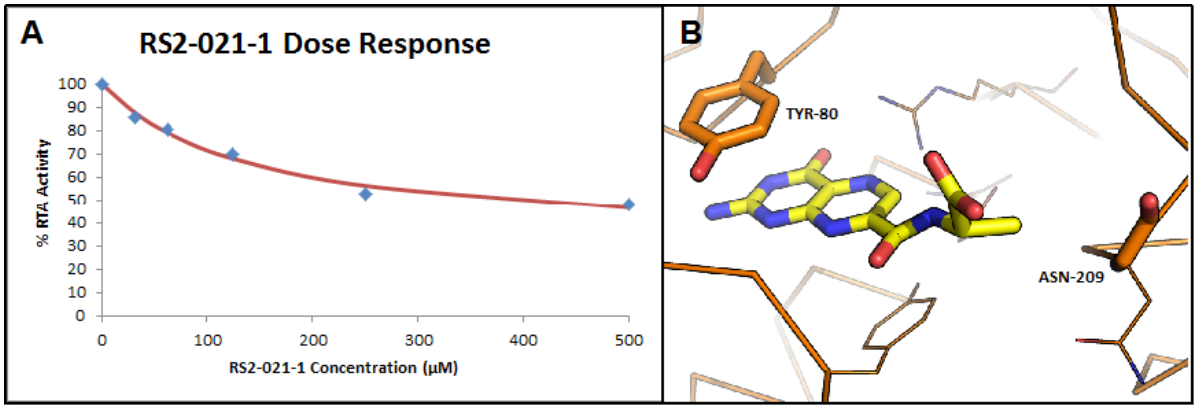

Figure 5: **A.** Dose response of RS2-021-1 averaged from six repetitions of the firefly luciferase assay. The IC_{50} was calculated as 350 μM. **B.** Complex crystal structure of RS2-021-1 (yellow) and RTA (orange). The compound forms no hydrogen bonds outside the primary binding pocket, but does make a hydrophobic contact with Asn-209.

The serine derivative RS2-022-1 showed a slight increase in potency over 7CP with an apparent IC_{50} of 175 μM. The crystal structure, which was solved in the P41212 space group at a resolution of 1.44 Å, is shown in Figure 6B. The serine side chain's hydroxyl group donates a hydrogen bond to the backbone carbonyl oxygen of Glu 208. The second pocket also appears to be occupied by another RS2-022-1 molecule, although the electron density was only enough to account for 50% occupancy. The tautomer of the pterin group of the second molecule with the proton on N3 is favored so that N1 and N8 can both accept hydrogen bonds from the side chain amide group of Asn 78. Additionally, the exocyclic amine donates a hydrogen bond to the Asp 75 side chain, and the 4-oxo atom accepts a hydrogen bond from the backbone NH of Asp 96. The OH group from the serine pendant also accepts a hydrogen bond from the guanidinium group of Arg 258. Because of the 50% occupancy, Tyr 80 has two distinct positions in the presence and absence of the second ligand, its inward facing conformation lining up with the position of the second pterin ring for a more favorable pi stacking interaction.

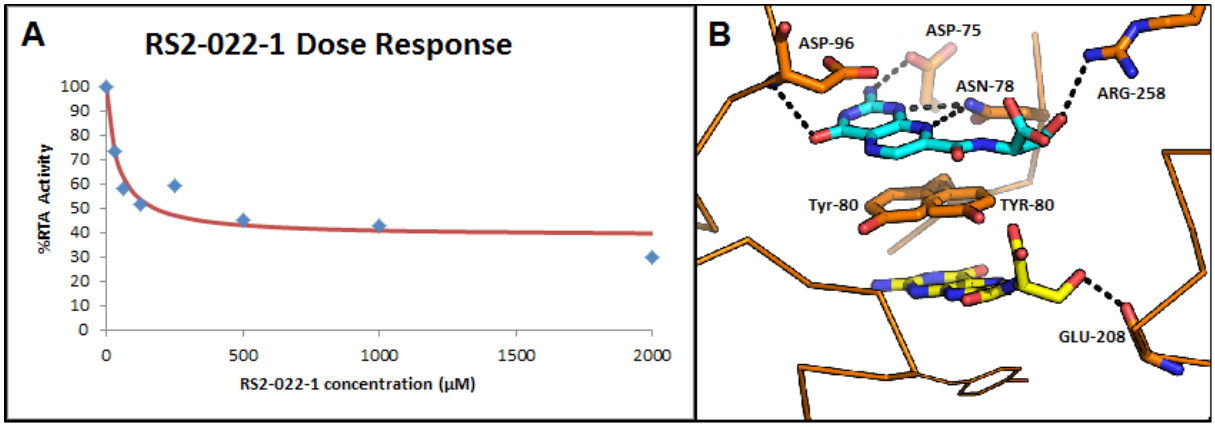

Figure 6: **A.** Dose response of RS2-022-1 averaged from six repetitions of the firefly luciferase assay. The IC_{50} was calculated as 175 µM. **B.** Complex crystal structure of RS2-022-1 (yellow) and RTA (orange). The compound forms a hydrogen bond (black dashed lines) with Glu-208. A second molecule of RS2-022-1 (blue) binds in the secondary pocket, making five hydrogen bonds with RTA.

6.3 – Dipeptide compounds

The three dipeptide derivative presented here, 7CP-Gly-Phe (RS2-058-1), 7CP-Gly-Tyr (RS2-136-1), and 7CP-Ser-Phe (RS2-150-1), are representative of early efforts to optimize the amino acid sequence of these inhibitors by varying the first and second amino acids independently. RS2-058-1, the first dipeptide tested had an apparent IC_{50} of 20 µM, a more than ten-fold improvement over 7CP. From the structure in Figure 7B, which was solved in the P41212 space group at a resolution of 1.52 Å, a unique edge to face stacking interaction between the edge of the Trp 211 and the face of the phenylalanine pendant is observed. This interaction alone could account for the increased binding affinity because it is roughly equivalent to the strength of a hydrogen bond [42, 43] and the water displaced for this interaction is energetically favorable since both structures are hydrophobic. Aromatic residues in the second position are expected to retain this favorable interaction.

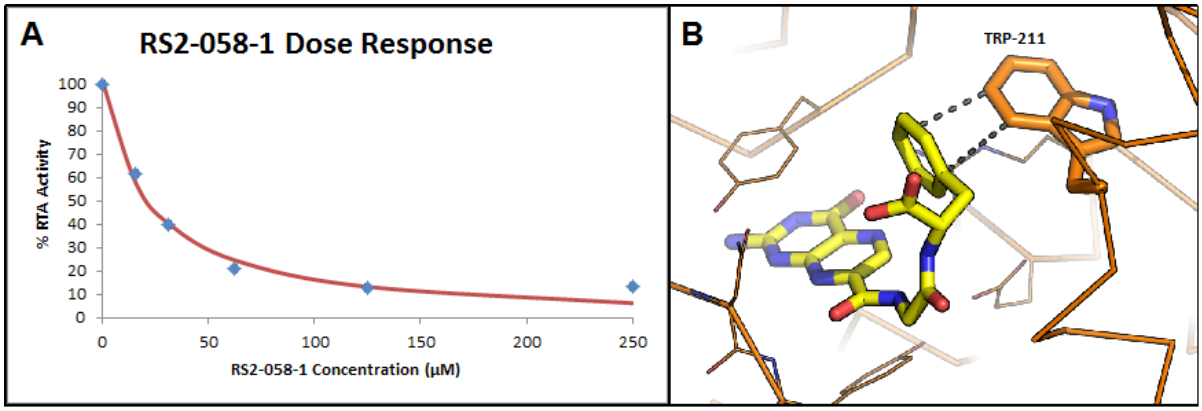

Figure 7: **A.** Dose response of RS2-058-1 averaged from six repetitions of the firefly luciferase assay. The IC_{50} was calculated as 20 μM. **B.** Complex crystal structure of RS2-058-1 (yellow) and RTA (orange). The phenyl group makes an edge-to-face stacking interaction (grey dashed lines) with Trp-211.

The 7CP-Gly-Tyr compound (RS2-136-1) had an IC_{50} of 15 μM, a slight improvement over the 7CP-Gly-Phe compound (RS2-058-1). As predicted, the crystal structure (Figure 8B), solved in the P41212 space group at a resolution of 1.58 Å, shows that this compound adopts an almost identical binding pose to RS2-058-1, with the aromatic tyrosine side chain in position for the edge to face stacking interaction with Trp 211. Contrary to what was expected, the phenolic hydroxyl group on the tyrosine is not involved in any hydrogen bonding. Instead, it is likely that the increased binding affinity of this compound over the phenylalanine-containing one is due to the electron donating character of the hydroxyl group, which enhances the edge to face stacking interaction by increasing the electron density of the pi system in the face.

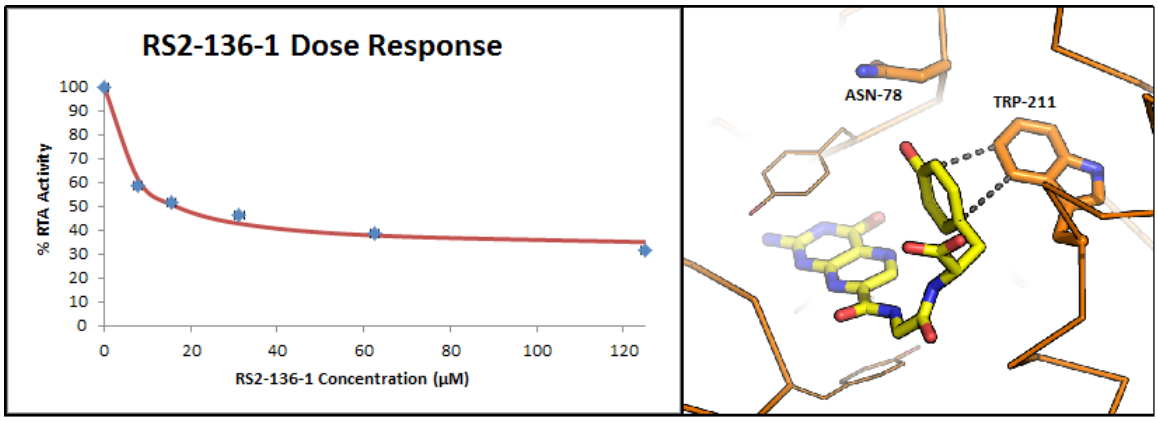

Figure 8: **A.** Dose response of RS2-136-1 averaged from six repetitions of the firefly luciferase assay. The IC_{50} was calculated as 15 µM. **B.** Complex crystal structure of RS2-136-1 (yellow) and RTA (orange). The phenol group makes an edge-to-face stacking interaction (grey dashed lines) with Trp-211.

Substitution of the glycine residue for serine in the first position to yield 7CP-Ser-Phe (RS2-150-1) led to a decrease in binding affinity with an apparent IC_{50} of 50 µM. The reason for this becomes quite apparent when looking at the crystal structure (Figure 9B), which was solved in the P41212 space group at a resolution of 1.75 Å. The edge to face interaction seen in the previous two structures is lost. Instead, Tyr 80 rotates out to accept a hydrogen bond from the NH of the amide linking the serine and phenylalanine residues of the ligand. The positioning of Tyr 80 prevents the phenylalanine of RS2-150-1 from orienting correctly for the interaction with Trp 211. There is another hydrogen bond made by the serine residue but it is donating to a malonate molecule rather than to a part of RTA. It is likely, however, that in normal conditions this hydrogen bond is donated to Glu 208 as was seen in the RS2-022-1 structure. From these results, we can see that the compactness and flexibility of glycine in the first position allows the second aromatic residue to make the edge to face stacking interaction with Trp 211.

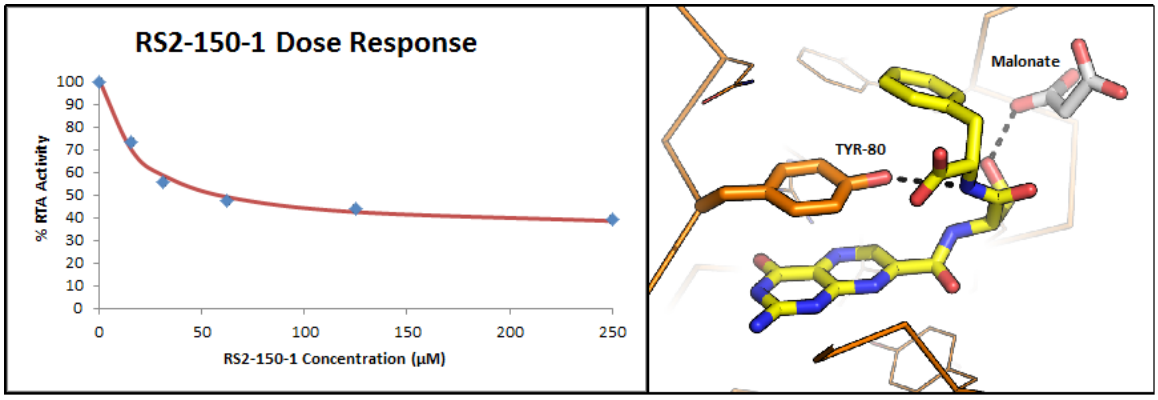

Figure 9: **A.** Dose response of RS2-150-1 averaged from six repetitions of the firefly luciferase assay. The IC_{50} was calculated as 50 µM. **B.** Complex crystal structure of RS2-150-1 (yellow) and RTA (orange). The compound forms two hydrogen bonds, shown as black dashed lines; one with Tyr-80 and the other with a malonate molecule (white).

6.4 – Tripeptide compounds

The first tripeptide compound tested, 7CP-Gly-Phe-Phe (RS2-123-2), was an extension of the potent inhibitor RS2-058-1 with another phenylalanine residue in the third position. What resulted was a moderate increase in binding affinity with RS2-123-2 having an apparent IC_{50} of 15 µM. The crystal structure (Figure 10B) was solved in the P41212 space group at a resolution of 1.7 Å, and shows that the phenylalanine in the second position makes the edge to face interaction with Trp 211 as expected. The extra phenylalanine residue makes no additional interactions with the protein, but is instead aligned parallel to the other phenylalanine. The two phenylalanine residues make an intramolecular pi stacking interaction which could possibly contribute to the increased potency by stabilizing the binding conformation.

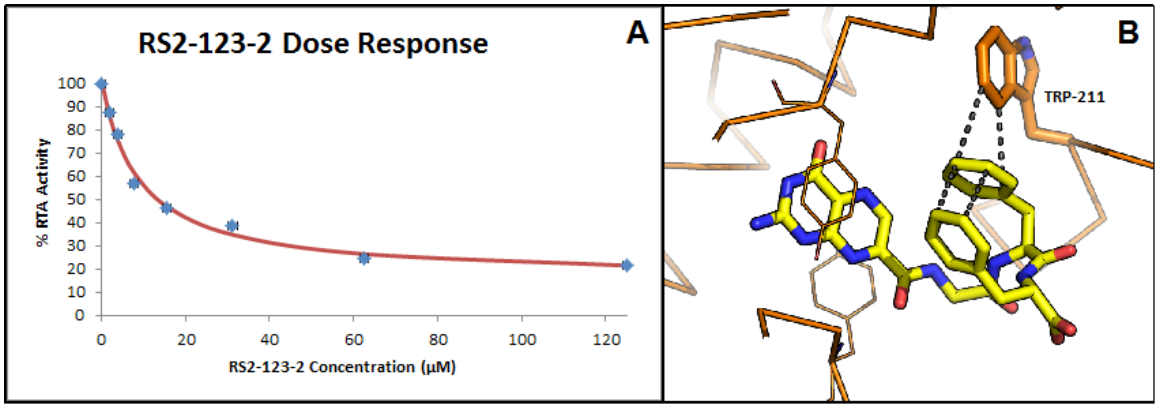

Figure 10: **A.** Dose response of RS2-123-2 averaged from six repetitions of the firefly luciferase assay. The IC_{50} was calculated as 15 μM. **B.** Complex crystal structure of RS2-123-2 (yellow) and RTA (orange). The two phenyl rings stack parallel with one another, while the first ring makes an edge-to-face stacking interaction (grey dashed lines) with Trp-211.

A virtual drug screen of 7CP tripeptides predicted that 7CP-Ser-Gly-Trp (RS2-149-1) would have the strongest binding affinity of the group. The resulting IC_{50} of this compound was 25 μM, making this a potent RTA inhibitor on par with RS2-058-1. The binding pose, particularly regarding the position of the tryptophan residue, is somewhat uncertain based on the ambiguous electron density at this position. As seen in the crystal structure shown in Figure 11B, which was solved in the P41212 space group at a resolution of 1.75 Å, the tryptophan of RS2-149-1 either makes a cation-pi interaction with Arg 258 and hydrophobic interactions with the residues in the surrounding surface, or makes the edge-to-face stacking interaction with Trp 211 while the carboxy terminus interacts with Arg 258. Additionally, the serine in the first position donates a hydrogen bond to the backbone carbonyl oxygen of Glu 208. This interaction is nearly identical to the one seen in the crystal structure of RS2-022-1, providing evidence that one should expect that this hydrogen bond be conserved in amino acid derivatives with serine at the first position.

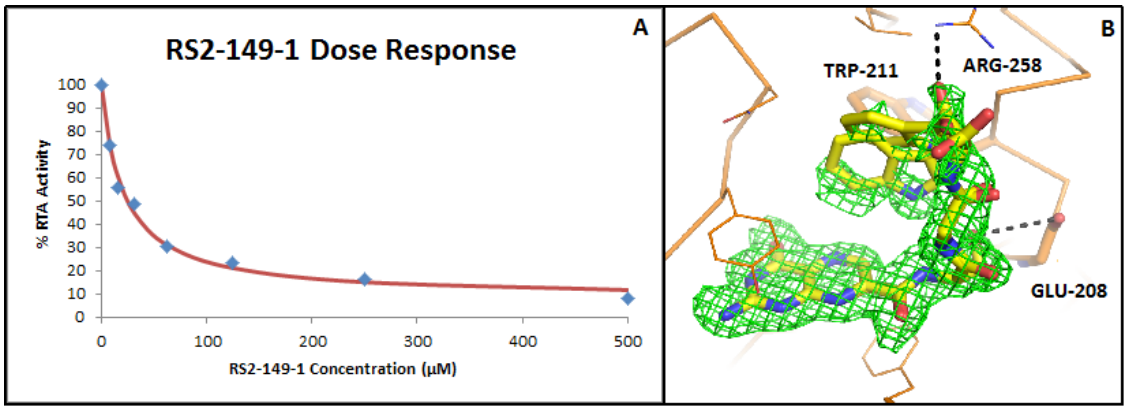

Figure 11: **A.** Dose response of RS2-149-1 averaged from six repetitions of the firefly luciferase assay. The IC_{50} was calculated as 25 μM. **B.** Complex crystal structure of RS2-149-1 (yellow) and RTA (orange) with Fo-Fc omit map (green) contoured at 2 sigma around the ligand. The compound forms a hydrogen bond, shown as black dashed lines, with Glu-208. The density for the indole ring of the compound suggests that multiple binding conformations exist: one in which the indole forms a cation-pi interaction with Arg-258 and makes hydrophobic contacts with the surface around this residue, and another in which the indole makes an edge-to-face stacking interaction with Trp-211 while the carboxylic acid forms an ionic/hydrogen bond with Arg-258.

Chapter 7: Discussion

The application of structure-based drug design to the pterin-amino acid series of RTA inhibitors has provided a wealth of valuable information. The inhibitory data and complex structures for each of these compounds give some insight into a rough structure-activity relationship. From the data for the series of dipeptide and tripeptide 7CP derivatives with aromatic amino acids in the second position, it is evident that the edge to face stacking interaction observed in these structures contributes greatly to activity. For all compounds that make this interaction, the IC_{50} was at least ten times lower than that of 7CP. Of the compounds lacking this feature, none performed as well in the assay, although RS2-150-1 and RS2-149-1 came close with IC_{50} values of 50 μM and 25 μM respectively. In each of these structures, multiple ligand-receptor interactions were observed in the amino acid pendants which contributed to the increased binding affinity.

7.1 – Ideal amino acid sequences for further extensions

The consistency observed in the structures allows for predictions to be made regarding the binding pose and performance of future compounds based on the amino acid sequence of their pendants. For example, a serine in the first position will likely donate a hydrogen bond to Glu 208, and an aromatic second will probably make an edge to face pi stacking interaction with Trp 211. From the structure of RS2-150-1, however, we know that these two features are unable to be combined in the same molecule, meaning that glycine must be used in the first position in order for the aromatic residue in the send position to make its interaction. We also know now that the ideal aromatic amino acid for this interaction is tyrosine.

Due to the failure of any of the tripeptide compounds to extend into the second pocket, it is apparent that at least a four amino acid chain would be needed for this task.

From the limited data available for tripeptides, it appears as though the third amino acid position is of little importance if the second one is aromatic. This is due to the fact that the third amino acid is in an isolated position too far away from the protein surfaces to make any useful contacts. Knowing this and the ideal first two amino acids, the logical starting point for designing four amino acid derivatives of 7CP would begin with the sequence Gly-Tyr-Gly with an aromatic residue such as tryptophan in the fourth position which could potentially make a stacking interaction with Tyr 80 in the second pocket. The glycine in the third position, similar in role to the one in the first position, could provide the flexibility needed for the aromatic residue in the next position to extend into the second pocket.

Another route worth exploring would be 3-4 amino acid derivatives with serine at the first two positions. There are several potential hydrogen bond donors and acceptors in the vicinity that could form one or more contacts with the second serine, which could possibly make up for the lack of an edge to face stacking interaction provided by aromatic amino acids. Depending on how the peptide chain orients, an aromatic amino acid in the third position or one in the fourth preceded by a glycine could potentially stack in the second pocket.

7.2 – Occupation of the second pocket

The crystal structure for RS2-022-1 in complex with RTA had a unique feature that has never been seen before in RTA complex structures. The pterin appears in both the primary and secondary binding pocket, providing valuable insight into how pendants should be designed to promote binding in the second pocket. From the structure, it appears that the two molecules are bound close enough together to be connected together by a short linker such as ethylene diamine. It would be worthwhile to have such a

symmetric bis-7CP-Ser compound synthesized to have analyzed to see if it is indeed possible to have both pockets occupied by a single bidentate ligand.

Pterins have notorious solubility problems, so it is likely that adding a second pterin ring to the inhibitor platform will be problematic. Fortunately, the binding in the second pocket is not as specific to the pterin as the primary pocket is as only 4 hydrogen bonds are made. From the structure, we know that N1, N8, and O4 act as hydrogen bond acceptors and the exocyclic amine donates one hydrogen bond. The other two hetero atoms of the pterin do not appear to be needed for binding specificity. This in in agreement with how the guanine base binds the second pocket in the published tetranucleotide substrate analog structure [10], as all of the same amino acids are involved in hydrogen bonding. Also, the interaction with Arg 258 appears to be an important anchor, as this residue makes an ion pair with the phosphate in the tetranucleotide structure. In light of these similarities, an obvious alternative to the pterin ring in the second pocket would be guanine substituted at the 8 position, which would be analogous to the 7 position on the pterin.

7.3 – Peptide alternatives

While this approach of using peptide pendants has yielded potent RTA inhibitors, they are unlikely to be effective in an *in vivo* system due to degradation by proteases. It is therefore necessary, once ideal peptide pendants have been identified, to eliminate peptide bonds from the structure while maintaining the important features of the peptide inhibitor. There are many isosteres for peptides but they often expand or restrict the flexibility of the parent structure, so it is unpredictable how these potential analogs will perform. An example of such analogs is a series of RTA inhibitors that are based on the triazole linker derived from copper catalyzed click chemistry [44]. Compounds bearing

pendants in which the triazole is followed by an aromatic group perform well in the luciferase assay and bind in a similar fashion to the glycine aromatic dipeptide compounds. Despite making the same interactions as these dipeptides, the triazoles lose some flexibility making them too conformationally restrained to position the aromatic group at the ideal distance and angle for the edge to face interaction with Trp 211, so their inhibition of RTA is not as strong. Triazoles and other potential linkers such as hydrazides, sulfonamides, and ureas need to be explored with the goal of mirroring the features of the peptide pendants. Although these peptide based RTA inhibitors are unlikely to become effective drugs, they will provide a wealth of structural information to be used as a foundation for the design of RTA drugs.

References

1. Olsnes, S.; Pihl, A., Toxic Lectins and Related Proteins. In *The Molecular Action of Toxins and Viruses*, Coen, P.; van Heynigen, S., Eds. Elsevier Biomedical Press: New York, 1982.
2. Stirpe, F.; Barbieri, L., Ribosome-inactivating proteins up to date. *FEBS Lett* **1986,** 195, (1-2), 1-8.
3. Ready, M. P.; Brown, D. T.; Robertus, J. D., Extracellular localization of pokeweed antiviral protein. *Proc Natl Acad Sci U S A* **1986,** 83, (14), 5053-6.
4. Burnham, R. M., Threat of Bioterrorism in America. In Subcommittee on Oversight and Investigations: Report before the U.S. House of Representatives, 1999.
5. Rich, V., Murderous Experiments of Stalin's Police Chief. *New Scientist* 1992, p 8.
6. Mishra, V.; Bilgrami, S.; Sharma, R. S.; Kaur, P.; Yadav, S.; Krauspenhaar, R.; Betzel, C.; Voelter, W.; Babu, C. R.; Singh, T. P., Crystal structure of himalayan mistletoe ribosome-inactivating protein reveals the presence of a natural inhibitor and a new functionally active sugar-binding site. *J Biol Chem* **2005,** 280, (21), 20712-21.
7. Montfort, W.; Villafranca, J. E.; Monzingo, A. F.; Ernst, S. R.; Katzin, B.; Rutenber, E.; Xuong, N. H.; Hamlin, R.; Robertus, J. D., The three-dimensional structure of ricin at 2.8 A. *J Biol Chem* **1987,** 262, (11), 5398-403.
8. Rutenber, E.; Katzin, B. J.; Ernst, S.; Collins, E. J.; Mlsna, D.; Ready, M. P.; Robertus, J. D., Crystallographic refinement of ricin to 2.5 A. *Proteins* **1991,** 10, (3), 240-50.
9. Monzingo, A. F.; Robertus, J. D., X-ray analysis of substrate analogs in the ricin A-chain active site. *J Mol Biol* **1992,** 227, (4), 1136-45.
10. Ho, M. C.; Sturm, M. B.; Almo, S. C.; Schramm, V. L., Transition state analogues in structures of ricin and saporin ribosome-inactivating proteins. *Proc Natl Acad Sci U S A* **2009,** 106, (48), 20276-81.
11. Katzin, B. J.; Collins, E. J.; Robertus, J. D., Structure of ricin A-chain at 2.5 A. *Proteins* **1991,** 10, (3), 251-9.
12. Weston, S. A.; Tucker, A. D.; Thatcher, D. R.; Derbyshire, D. J.; Pauptit, R. A., X-ray structure of recombinant ricin A-chain at 1.8 A resolution. *J Mol Biol* **1994,** 244, (4), 410-22.
13. Robertus, J. D.; Monzingo, A. F., The structure of ribosome inactivating proteins. *Mini Rev Med Chem* **2004,** 4, (5), 477-86.
14. Mlsna, D.; Monzingo, A. F.; Katzin, B. J.; Ernst, S.; Robertus, J. D., Structure of recombinant ricin A chain at 2.3 A. *Protein Sci* **1993,** 2, (3), 429-35.
15. Hajduk, P. J.; Huth, J. R.; Fesik, S. W., Druggability indices for protein targets derived from NMR-based screening data. *J Med Chem* **2005,** 48, (7), 2518-25.

16. Lord, J. M.; Deeks, E.; Marsden, C. J.; Moore, K.; Pateman, C.; Smith, D. C.; Spooner, R. A.; Watson, P.; Roberts, L. M., Retrograde transport of toxins across the endoplasmic reticulum membrane. *Biochem Soc Trans* **2003,** 31, (Pt 6), 1260-2.
17. Sandvig, K.; Olsnes, S.; Pihl, A., Kinetics of binding of the toxic lectins abrin and ricin to surface receptors of human cells. *J Biol Chem* **1976,** 251, (13), 3977-84.
18. Endo, Y.; Tsurugi, K., RNA N-glycosidase activity of ricin A-chain. Mechanism of action of the toxic lectin ricin on eukaryotic ribosomes. *J Biol Chem* **1987,** 262, (17), 8128-30.
19. Endo, Y.; Tsurugi, K., The RNA N-glycosidase activity of ricin A-chain. The characteristics of the enzymatic activity of ricin A-chain with ribosomes and with rRNA. *J Biol Chem* **1988,** 263, (18), 8735-9.
20. Olsnes, S.; Fernandez-Puentes, C.; Carrasco, L.; Vazquez, D., Ribosome inactivation by the toxic lectins abrin and ricin. Kinetics of the enzymic activity of the toxin A-chains. *Eur J Biochem* **1975,** 60, (1), 281-8.
21. Ready, M. P.; Kim, Y.; Robertus, J. D., Site-directed mutagenesis of ricin A-chain and implications for the mechanism of action. *Proteins* **1991,** 10, (3), 270-8.
22. Chen, X. Y.; Link, T. M.; Schramm, V. L., Ricin A-chain: kinetics, mechanism, and RNA stem-loop inhibitors. *Biochemistry* **1998,** 37, (33), 11605-13.
23. Cheng, A. C.; Coleman, R. G.; Smyth, K. T.; Cao, Q.; Soulard, P.; Caffrey, D. R.; Salzberg, A. C.; Huang, E. S., Structure-based maximal affinity model predicts small-molecule druggability. *Nat Biotechnol* **2007,** 25, (1), 71-5.
24. Wahome, P. G.; Bai, Y.; Neal, L. M.; Robertus, J. D.; Mantis, N. J., Identification of small-molecule inhibitors of ricin and shiga toxin using a cell-based high-throughput screen. *Toxicon* **2010**.
25. Stechmann, B.; Bai, S. K.; Gobbo, E.; Lopez, R.; Merer, G.; Pinchard, S.; Panigai, L.; Tenza, D.; Raposo, G.; Beaumelle, B.; Sauvaire, D.; Gillet, D.; Johannes, L.; Barbier, J., Inhibition of retrograde transport protects mice from lethal ricin challenge. *Cell* **2010,** 141, (2), 231-42.
26. Roday, S.; Amukele, T.; Evans, G. B.; Tyler, P. C.; Furneaux, R. H.; Schramm, V. L., Inhibition of ricin A-chain with pyrrolidine mimics of the oxacarbenium ion transition state. *Biochemistry* **2004,** 43, (17), 4923-33.
27. Bai, Y.; Monzingo, A. F.; Robertus, J. D., The X-ray structure of ricin A chain with a novel inhibitor. *Arch Biochem Biophys* **2009,** 483, 23-28.
28. Miller, D. J.; Ravikumar, K.; Shen, H.; Suh, J. K.; Kerwin, S. M.; Robertus, J. D., Structure-based design and characterization of novel platforms for ricin and shiga toxin inhibition. *J Med Chem* **2002,** 45, (1), 90-8.
29. Yan, X.; Hollis, T.; Svinth, M.; Day, P.; Monzingo, A. F.; Milne, G. W.; Robertus, J. D., Structure-based identification of a ricin inhibitor. *J Mol Biol* **1997,** 266, (5), 1043-9.
30. Chang, C. E.; Chen, W.; Gilson, M. K., Ligand configurational entropy and protein binding. *Proc Natl Acad Sci U S A* **2007,** 104, (5), 1534-9.

31. Waring, P.; Armarego, W. L. F., Pterins .10. A New Preparation of 6-Hydroxymethylpterin from 6-Methylpterin. *Australian Journal of Chemistry* **1985,** 38, (4), 629-631.
32. Taylor, E. C.; Kobylecki, R., Pteridines .43. Facile Synthesis of 6-Chloropterin and 2,4-Diamino-6-Chloropteridine. *Journal of Organic Chemistry* **1978,** 43, (4), 680-683.
33. Taylor, E. C.; Jacobi, P. A., Letter: An unequivocal total synthesis of L-erythro-biopterin. *Journal of the American Chemical Society* **1974,** 96, (21), 6781-2.
34. Taylor, E. C.; Henrie, R. N.; Portnoy, R. C., Pteridines .44. Convenient Synthesis of 6-Formylpterin. *Journal of Organic Chemistry* **1978,** 43, (4), 736-737.
35. Pruet, J. M.; Robertus, J. D.; Anslyn, E. V., Acyl radical insertion for the direct formation of new 7-substituted pterin analogs. *Tetrahedron letters* **2010,** 51, (18), 2539-2540.
36. Pruet, J. M.; Jasheway, K. R.; Manzano, L. A.; Bai, Y.; Anslyn, E. V.; Robertus, J. D., 7-Substituted pterins provide a new direction for ricin A chain inhibitors. *European Journal of Medicinal Chemistry* **2011,** In Press, Corrected Proof.
37. Feng, B. Y., Shelat, A., Doman, T., Guy, R. K., Shoichet, B. K., High-throughput assays for promiscuous inhibitors. *Nature Chemical Biology* **2005** 1, 146-8.
38. Otwinowski, Z. and W. Minor. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **1997,** 27, 307-26.
39. Vagin, A. and A. Teplyakov. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **1997,** 30, 1022-5.
40. Murshudov, G. N., A. A. Vagin and E. J. Dodson. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **1997,** 53, (3), 240-55.
41. A. W. Schüttelkopf and D. M. F. van Aalten. PRODRG - a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D* **2004,** 60, 1355-63.
42. Kim, E., Paliwal, S., Wilcox, C. S., Measurements of Molecular Electrostatic Field Effects in Edge-to-Face Aromatic Interactions and CH-π Interactions with Implications for Protein Folding and Molecular Recognition. *Journal of the American Chemical Society* **1998** 120, (43), 11192-3.
43. Jennings, W. B., Farrel, B. M., Malone, J. F., Attractive Intramolecular Edge-to-Face Aromatic Interactions in Flexible Organic Molecules. *Accounts of Chemical Research* **2001** 34, (11), 885-94.
44. Rostovtsev, V. V., Green, L. G., Fokin, V. V., Sharpless, K. B., A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angewandte Chemie* **2002,** 114, (14), 2708-11.